

dunetpc - Task #24595

Add dataprep for simulation reco

07/03/2020 10:09 AM - David Adams

Status:	Closed	Start date:	07/03/2020
Priority:	Normal	Due date:	
Assignee:	Tingjun Yang	% Done:	100%
Category:		Estimated time:	0.00 hour
Target version:			
Description			
At present simulation reco is run without the dataprep stage. We would like to add this.			

History

#1 - 07/03/2020 10:22 AM - David Adams

First step is to create a dataprep configuration and so this is assigned to me for now.

#2 - 07/03/2020 11:24 AM - David Adams

What is the top-level fcl we use for reco of simulated data?

#3 - 07/03/2020 11:29 AM - Tingjun Yang

https://wiki.dunescience.org/w/img_auth.php/f0/G4_stage_split.png

It is protoDUNE_refactored_reco.fcl.

#4 - 07/03/2020 03:01 PM - David Adams

- Assignee changed from David Adams to Tingjun Yang

I have pushed some sequences for dataprep of simulated data. Tingjun, can you add this or its equivalent to the fcl preamble:

```
physics.producers.dataprep:@local::producer_adcprep_notool
physics.producers.dataprep.DigitLabel: "tpcrawdecoder:daq"
physics.producers.dataprep.WireName: "dataprep"
services.RawDigitPrepService.AdcChannelToolNames: @local::pdsim_dataprep_tools_wirecell
```

Then adding dataprep to the producers should produce the input for wirecell and Wenqiang can make the change to use that container as input.

The scale won't be quite right due to differences between nominal and the various other calibration scales. We can straighten that out next week.

I would like to present my part of this at the DRA meeting next week. There are issues about which tail removal and what other tools to include in the sequence.

#5 - 08/18/2020 11:03 PM - Tingjun Yang

I have started looking into this. I prefer to disable the tail simulation and reconstruction since the current tail simulation is different from data.

The pdsim_dataprep_tools_wirecell contains two tools:

```
adcScaleAdcToKe: {
  InputUnit: "ADC count"
  LogLevel: 1
  OutputUnit: "ke/tick"
  ScaleFactor: 2.401e-2
  tool_type: "AdcSampleScaler"
}
adcScaleKeToAdc: {
  InputUnit: "ke/tick"
  LogLevel: 1
  OutputUnit: "ADC count"
  ScaleFactor: 4.1649e1
  tool_type: "AdcSampleScaler"
}
```

The first tool converts ADC to Ke and the second tool converts Ke to ADC. Since we do not simulate gain variations between channels, the scale

factors of the two tools cancel each other, so this dataprep step simply changes the raw ADC from the format of raw::RawDigit to recob::Wire. We can still do this to make the MC reconstruction more consistent with data, but I just want to make sure people are aware of this.

Wenqiang, currently wclsdatanfsp in the MC reconstruction takes raw::RawDigit as input. Can it be configured to use recob::Wire as input?

#6 - 08/19/2020 08:53 AM - David Adams

In dataprep, we apply calibration early on and most of the work is done at ke scale. This removes the channel-to-channel non-uniformities and also allows noise intermediate noise measurements, event displays etc at that scale. For simulation, the initial scale change allows us to use the same tool configurations that are used for data.

IMHO, we should not apply the second scale change. We have calibrated the signals and should write it out at charge scale. Wirecell or any other client can apply any factor they wish after reading the dataprep output. Note that with this convention, one only needs to measure the area of a collection pulse to get the collected charge in ke. We can apply a factor if we wish to use a different charge unit, e.g. e or 200 e but converting back to average ADC response is confusing especially as our calibration evolves and that value changes.

As for the tail removal (and similar comments apply for noise), I think we should include it in the simulation and make corrections because there are cases where the correction fails, some expected and likely some surprises as well. The current simulation appears to have the wrong tail and should be fixed.

#7 - 08/19/2020 09:27 PM - Wenqiang Gu

Wenqiang, currently wclsdatanfsp in the MC reconstruction takes raw::RawDigit as input. Can it be configured to use recob::Wire as input?

In the `dunetpc/fcl/protodune/reco/protoDUNE_refactored_reco.fcl`, the signal processing for recob::Wire will work if you change the line

wclsdatanfsp: @local::protodunespdata_nfsp

to

wclsdatanfsp: @local::protodunespdata_wctsp

By the way, the signal processing for data assumes a few bad channels, do we want to "disable" these channels in the simulation?

#8 - 08/20/2020 04:08 PM - Tingjun Yang

Thanks Wenqiang for the instructions. I added dataprep and configured wirecell to use output from dataprep as input. Currently we zero signals from bad channels in the simulation. I think we should continue to do so.

I also changed the default bad channel list to pdsp_channel_status_2018 since we will reprocess the beam data.

#9 - 08/21/2020 07:24 AM - David Adams

I don't know what you mean by "changed the default bad channel list to pdsp_channel_status_2018" but this sounds like something that should be done in the high-level production-specific configuration. --da

#10 - 10/25/2020 09:31 PM - Tingjun Yang

The tail simulation is enabled with the correct configuration (1 layer): [ca2c542f2093f29a92ae1e374f50da115429ec0c](#).

The tail removal in dataprep for MC is also enabled: [3775d7b240b5e88e8fe68b94e0fe4753af5f5779](#).

#11 - 01/15/2021 04:31 PM - David Adams

Is this task complete?

#12 - 01/15/2021 04:39 PM - Tingjun Yang

- % Done changed from 0 to 100

- Status changed from Assigned to Resolved

#13 - 01/15/2021 04:39 PM - Tingjun Yang

- Status changed from Resolved to Closed